AD_____

GRANT NUMBER DAMD17-97-1-7316

TITLE: Development and Use of Novel Tools to Directly Screen for Substrates of Cyclin Dependent Kinases

PRINCIPAL INVESTIGATOR: Gavin J. Sherlock, Ph.D.

CONTRACTING ORGANIZATION: Cold Spring Harbor Laboratory
Cold Spring Harbor, New York 11724

REPORT DATE: August 1998

TYPE OF REPORT: Annual

PREPARED FOR: Commanding General
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1998	Annual (1 Jul 97 - 30 Jun 98)

4. TITLE AND SUBTITLE
Development and Use of Novel Tools to Directly Screen for Substrates of Cyclin Dependent Kinases

5. FUNDING NUMBERS
DAMD17-97-1-7316

6. AUTHOR(S)
Gavin J. Sherlock, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York 11724

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

19990301011

13. ABSTRACT *(Maximum 200 words)*

Cell division is a central and primary biological function. Many problems in biology and medicine center around cell division, but without a knowledge of the mechanism by which division is controlled, these problems are difficult to attack. The major transitions of the cell cycle are is contolled by a protein called a cyclin dependent kinase, or Cdk. The cell cycles of both yeast and human cells are controlled by a Cdk. Using the yeast two hybrid system I have isolated many novel Cdk interacting proteins in yeast, and these may be instrumental in controlling cell cycle progression. I have also isolate two exciting proetins which can only interact with a substrate trapping mutant of a yeast Cdk. These may be Cdk new substrates.

I have also used the genomic technology of "Gene Chips", to define the complete set of cell cycle regulated genes in yeast, and to characterise their contolling promoter elements.

These two projects have yielded novel and exciting results, of interest to both the yeast and cell cycle fields.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
24

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

 In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

 In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

 In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

8-9-98

PI - Signature Date

Table of Contents:

Front Cover 1

Standard Form 2

Foreword 3

Table of Contents: 4

Introduction: 5

Body: 8

Positives of directed two-hybrid screen 10

Known Cyclins: 10

Other known interactors: 10

New Interactors: 10

Genome wide analysis of cyclin specificity and cell cycle transcription: 12

Methods: 13

Conclusions: 17

References: 18

Appendices: 20

Curriculum vitae: 20

Introduction:

Cell division is a central and primary biological function. Many problems in biology and medicine center around cell division--cancer, aging, development, and regeneration, to name a few--but without a knowledge of the mechanism by which division is controlled, these problems are difficult to attack. In the budding yeast, *Saccharomyces cerevisiae*, as in all eukaryotes, the cell cycle can be divided into G1-, S-, G2- and M-phases. The major controlling event, termed Start, occurs late in G1, and once passed commits the cell to a round of division. Human cells have a G1 commitment event, called the restriction point, which is analogous, and probably homologous to Start. Before commitment, cells are influenced by signals such as pheromones, hormones, growth factors, nutrients, the rate of protein synthesis, and cell size. Once the choice for mitotic division has been made the cycle is usually completed, and is no longer greatly influenced by environmental signals. The overall goal of this proposal is to find out in molecular detail what the commitment event is.

In all eukaryotes, protein kinases called Cyclin Dependent Kinases (Cdks) regulate commitment (1). The two prototypic Cdks are Cdc28, from *S. cerevisiae*, and cdc2, from *S. pombe*. The Cdk is inactive as a monomer. It becomes active when it is appropriately phosphorylated, and when it associates with a regulatory subunit called a cyclin (2). The cyclin is necessary for kinase activity because its binding changes the structure of the Cdk in a way that allows access of substrates to the active site. In addition active site residues are realigned, allowing for a productive re-orientation of ATP (3). Hence Cdk binding of substrates is likely to be cyclin dependent.

In *S. cerevisiae* nine cyclins can bind and activate Cdc28, Cln1-Cln3 and Clb1-Clb6. Ten other cyclins, Pcl1, 2, 5-10, Clg1 and Pho80 bind to and activate a different Cdk, called Pho85

(4, 5). Pho85 is primarily involved in phosphate and glycogen metabolism, but under some conditions, when Cdc28-Cln functions are partially defective, Pcl-Pho85 can provide some overlapping functions (4, 6). Cln1, Cln2, and Cln3 promote G1 events, and so are called G1 cyclins (7, 8, 9). Pcl1 and Pcl2 can also affect G1 events. Clb5 and Clb6 are B-type cyclins which have roles in G_1 and S-phase (10, 11, 12). Their primary role is in the initiation of S-phase, though they are expressed in G1, and under some conditions they can substitute for Cln1 and Cln2. Clb1-Clb4, functionally, are mitotic cyclins: Clb3 and Clb4 help form the mitotic spindle in early mitosis, Clb1 and Clb2 are essential for elongation of the spindle, and other late mitotic events (13). Many of these cyclins are also important for controlling their own expression, or the expression of other cyclins.

Despite assembly of Cdc28-cyclin kinase complexes being necessary for the completion of START, no essential *in vivo* substrate for any of the complexes has been identified. *In vitro*, the different cyclin-Cdc28 complexes show no obvious specificity. In contrast, *in vivo*, the different cyclin-Cdc28 complexes promote different events. Cln3-Cdc28 is specific activator of the transcription factor SBF (9, 14, 15), whereas Clb2-Cdc28 inhibits SBF (16). Cln2-Cdc28 is good at phosphorylating Sic1 (17), at turning off signaling through the mating type response pathway (18), at promoting budding (19), and at shutting off the Anaphase Promoting Complex (APC) (19). Clb5 and other B-type cyclins cause initiation of DNA synthesis (12), while the Clns seem to be entirely incapable of this. Clb2 can promote mitosis, activate its own transcription, and shut off SBF, whereas the Clns are not capable of any of these functions even if drastically over-expressed (16, 20). Ectopic expression of cyclins from the *GAL* promoter shows that these differences are not just due to differences in time or amount of expression.

Thus, an important problem for mammalian as well as yeast cyclins is: What is the basis of the specificity of cyclin-Cdc28 complexes for different events? One idea is that the cyclins act as "targeting subunits" for the kinase, either helping bind directly to the substrate, or binding to other proteins to localize the kinase to certain structures or intra-cellular compartments. Of particular relevance is the fact that different endogenous proteins are found bound to different yeast cyclin-Cdc28 complexes (9, 21). This kind of specific binding may be the basis of cyclin specificity.

My original proposal sought to identify potential Cdk substrates, and the basis of substrate specificity with respect to the cyclin binding partner, by the use of two complimentary methods. Firstly by mutagenesis of the G1 cyclin CLN2, to create new alleles defective for specific functions, which could then be used as tools for the identification of substrates, and secondly using the two hybrid system to identify cyclin specific Cdc28 interacting proteins. Hence this would combine both genetic and biochemical approaches to help solve these important biological problems. At the time that this grant came into effect another group published a paper where they had carried out very similar work to that which I had proposed, with respect to the alanine scanning of Cln2 (22). Bearing this in mind, and not wanting to repeat published work, I decided to concentrate on my two-hybrid analysis idea. In addition, in collaboration with David Bostein's lab, I have been using the genomic technology of microarrays to investigate the functional differences between cyclins in their abilities to induce and repress transcription, and to define the complete set of cell cycle regulated genes in yeast. In both cases I feel that I have made important breakthroughs, which will contribute significantly to the fields of both yeast and cell cycle research. I describe these results below.

Body:

I had initially intended to screen for Cdc28 interacting proteins, whose interaction was Cln2 dependent. Because Fred Cross's group was concentrating several of their efforts on identifying substrates of Cln2-Cdc28 complexes (22, 23), I instead chose to screen for Clb5 dependent positives. However using a lexA-CDC28 fusion, in the presence of a plasmid overexpressing CLB5 from the galactose inducible promoter, I was only able to isolate clones encoding the known Cdc28 interacting protein Cks1, and this interaction was not dependent on Clb5. I therefore considered a second method for identifying Cdc28 substrates. I altered the ATP binding domain of Cdc28 to create a mutant that is inactive for kinase activity, yet should be able to bind both cyclins and ATP. It was envisioned that this mutant Cdc28 would still bind substrates, yet because it can no longer phosphorylate them, would not release them, hence strengthening the interaction enough to be detected by the two-hybrid system. This is similar in principle to a method used to isolate phosphatase substrates (24), which was very successful.

I constructed a fusion between the Gal4 activation domain, and either wild-type Cdc28, or Cdc28 in which I had mutated lysine 40 to an arginine. The comparable mutation in other protein kinases has been shown to result in a 'dead kinase'. I then screened for proteins that could interact with my mutated Cdc28 fusion, this time using a new library and strain (25). From over 3 million transformants I recovered only 12 positive clones, and only three of these were able to retransform to the selective phenotype. Sequencing revealed that one of these clones encoded the G1 cyclin Cln3, a known Cdc28 interactor. The other two clones both encoded different fusions of a protein Zds1. Remarkably both Zds1 clones could interact with the mutant Cdc28, but not with the wild-type, suggesting that the strength of this interaction was not sufficiently detectable when Cdc28 was able to function as a protein kinase. This strongly

suggests that Zds1 may be a Cdc28 substrate. One Zds1 fusion encompassed amino acids 529 through 915 (the end of the protein), the other, amino acids 569 through 802, suggesting these 233 amino acids are sufficient for the interaction. This region does not show significant homology to any other protein in the database, but contains 4 tandem copies of a novel repeat motif, VQxS(AV)P. It is not known whether this is significant. In addition this region contains a serine rich stretch of amino acids, which contains three tandem copies of the sequence SSSP. SP or TP maybe the minimum information required for Cdc28 phosphorylation. *ZDS1* was isolated as a high copy suppressor of the plasmid loss defect of a *cdc28*-1N allele, a defect that could not be suppressed by a close homologue of *ZDS1*, *ZDS2* (26). More recently it has been suggested (27) that Zds1 acts to transcriptionally repress *SWE1*, which encodes a tyrosine kinase that inhibits Cdc28 activity. A physical interaction between Zds1 and Cdc28 may therefore set up a negative feedback loop to regulate Cdc28 activity. Genetic results linking Zds1 to Cdc28 make the 2-hybrid interaction, especially one that is restricted to the dead-kinase version of Cdc28, appear compelling. To characterise this interaction further I intend to attempt to demonstrate that Zds1 can co-precipitate with at least mutant Cdc28. To this end I have tagged Zds1 with a triple HA epitope, so now have the immunological reagents necessary for the experiment.

In addition, because the number of isolated Cdc28-interacting positives was low, and because many expected Cdc28 interactors were not isolated, a new screen was done. Reasons that many known interactors were not isolated include a) the selection conditions were too stringent, and b) productive fusions for some proteins do not exist in the screening library. To counteract both of these possibilities I entered into collaboration with Stan Fields' laboratory. Dr. Fields, one of the originators of the two-hybrid system, is developing a library of all possible

yeast two-hybrid fusions, by cloning them one by one into the relevant vectors. His library should therefore achieve complete coverage when finished. In addition the screening procedure is semi-automated, and therefore a number of different selection stringencies can easily be tested. I recloned my wild type and mutant Cdc28 into new vectors and Dr. Fields, and Becky Drees in his lab, have carried out a directed screen against their almost complete library. These screens yielded some exciting results, which are detailed below:

Positives of directed two-hybrid screen

Known Cyclins:

CLN1 G1 cyclin.

CLN2 G1 cyclin.

CLN3 G1 cyclin.

CLB1 Mitotic cyclin.

Other known interactors:

CAK1 Cdc28 activating kinase, a known interactor.

New Interactors:

PCL7 Pcl7 is a cyclin normally associated with Pho85p. All cyclins have a common structure called the cyclin fold that consists of 5 alpha helices, and a cross complex formation of a Pho85 cyclin with Cdc28 is not unreasonable. Indeed in a two hybrid screen using Pho85, in which *PCL7* was originally isolated, Cln1, Ume3 and Ccl1 were all isolated as Pho85 interactors, all of which are considered partners for other Cdks (Cdc28, Ume5 and Kin28 respectively)(5). Whether this interaction is functionally relevant is however

unclear. If in the absence of Pho85 (which is not essential), a *PCL7* deletion results in additional phenotypes, then Pcl7 may have some Pho85 independent functions.

SAP155 Sap155 associates with Sit4p, a phosphatase involved in cell cycle regulation (28). It is also a phosphoprotein *in vivo*, and the complex it forms with Sit4 persists from late G1 to late M-phase. Despite much work on Sit4, its role in the cell cycle, and its connection to Cdc28, have not been elucidated. The interaction of Sap155 with Cdc28 may be the crux of this connection.

STB1 A protein that binds Sin3p, a transcriptional regulator. Sin3 contains 9 potential Cdc28 consensus phosphorylation sites, which out of over 6000 yeast proteins places it 19th on the list of most sites. In fact it has the 5th highest density of Cdc28 sites of all yeast proteins.

SPC42 Spindle pole body component. Spc42 has also been isolated in an unpublished 2-hybrid screen using Clb5 as the bait. In addition Dr. Futcher has done some GFP studies with Clb5 and Clb3, both of which showed discrete spots corresponding to the Spindle pole bodies. It is possible that Spc42 is a substrate itself, or that it acts to actually localize Cdc28 complexes to the SPB where they can phosphorylate/interact with other targets.

ARP8 Actin related protein. Nothing much is really known about ARP8. It has 3 potential Cdc28 phosphorylation sites, and of all the actin-related proteins in yeast, it has the largest number of insertions between the actin conserved blocks of homology.

YKR077W This protein is a good candidate for a Cdc28 substrate because it contains 7 potential Cdc28 phosphorylation sites. I also have unpublished evidence that shows that the expression of YKR077W peaks at the G1/S phase boundary, suggesting a cell cycle role. We also know, from unpublished data, that YKR077W is non-essential, and we

didn't observe any obvious phenotype in the deletant. There is however a weak homologue, YOR066W. What gives us a little more confidence that it may be a real homologue is the fact that yeast has several duplicated chromosomal blocks. One block contains YKR077w, and the corresponding duplicate block contains YOR066W in the corresponding position. YOR066w also has a large number of potential sites, and from our unpublished data we also know it is cell cycle regulated, though in this case it peaks at the M/G1 boundary.

YDR130C Novel. This is of special interest because it interacted specifically with the dead-kinase version of Cdc28, strongly suggesting that it may be a substrate. It contains 5 potential phosphorylation sites, and is predicted to form a coiled-coil structure, which is often indicative of structural proteins, either with a role in the SPB, or microtubule biology. In addition I have unpublished data that show that the transcript is regulated, and that it peaks in S/G2, suggestive of a role in the cell cycle.

YPL014W Novel.

YPL070W Novel.

YKL014C Novel.

Genome wide analysis of cyclin specificity and cell cycle transcription:

Several months ago Dr. Futcher and I became aware that a group at Stanford University, headed by David Botstein, were endeavoring to use microarray technology (see (29, 30)) to define the complete set of cell cycle regulated genes in yeast. We had previously approached a group headed by Pat Brown at Stanford to use the microarray technology to characterize cyclin specificity in terms of their ability to induce and repress transcription of other genes. We offered our considerable expertise in cell cycle studies to help Dr. Botstein's group to define all cell cycle

regulated genes. These data, in combination with our studies of cyclin specificity, provide invaluable information for all cell cycle studies. Knowing how cyclins induce transcription of other genes, what these other genes are, and when their expression in the cell cycle peaks provides us with a comprehensive picture of cause and effect with respect to cell cycle regulation.

Methods:

We did two experiments to classify genes by their mode of regulation to look at cyclin specificity. Many genes expressed in G1 phase are known to be induced by the related transcription factors SBF and MBF, which bind to SCB and MCB sites respectively. These two factors are activated by Cln3-Cdc28 in G1. Later, Clb2-Cdc28 inactivates SBF and MBF, and also induces transcription of a new set of genes. To extend our knowledge of the genes regulated in these ways, we used a strain with conditional Cln activity and conditional Cdc34 activity. The strain was arrested at the cdc34 block point at the G1/S boundary in the absence of Cln activity; then, while still at the cdc34 block point, Cln3 activity was restored. Samples were taken at different times after the restoration of Cln3. Thus, in this experiment, we could see all the transcripts affected directly (or somewhat indirectly) by the presence of Cln3 without allowing any cell cycle progression. Similarly, we used a conditional clb1, 2, 3, 4 strain and blocked it in mitosis using nocodazole in the absence of Clb1, 2, 3 or 4 activity. Clb2 activity was then restored, still in the presence of nocodazole, and samples taken.

To address the cell cycle profile of all transcripts we synchronized yeast cells by three different methods: centrifugal elutriation, alpha-factor block-release, or cdc15 block-release. Synchronized populations were monitored through the cell cycle, and samples were taken at regular intervals.

RNA from each experimental sample was used to make cDNA labeled with a red fluorophor. RNA from control samples (e.g. for the cell cycle experiments, asynchronous populations were used), was used to generate cDNA labeled with a green fluorophor. For each timepoint equal amounts of experimental and control cDNA were competitively hybridized to microarrays, which contain all 6000 or so ORFs. The ratio of red to green signal provides an expression level for each transcript relative to the control at each timepoint. Thus because for each time series the control is the same, the fluctuation of every transcript can be followed through each time series.

To identify which transcripts were periodic through the cell cycle we used Fourier transformation to identify how well the time series for each gene corresponded to a periodic function. Using this method we have identified between 500 and 800 genes that are cell cycle regulated, depending on the cut-off value that we use from the Fourier transformation. I have included a figure in the appendix that shows a color representation for all our cell cycle regulated genes from the cyclin induction experiments, and the elutriation and alpha-factor experiments.

An extensive search of the literature revealed that there are 97 genes that are currently published as being cell cycle regulated. We were able to identify 84 of these 97 using our data and our Fourier analysis method, which is an 87% success rate. Of the 13 that our methods did not identify 5 are induced by alpha-factor, so may have been missed because of an artificially high level at the beginning of the alpha-factor block-release experiment. In addition 5 of the false negatives were already flagged as not being adequately present on the chips because of a failed or poor PCR. Hence our success rate is closer to 91%, which we feel is very high indeed. A more worrisome point to consider is the rate at which we are identifying as cell cycle regulated genes which in fact are not. One way to address whether noise in the data could simply generate

false positives was to randomize the data either horizontally (that is randomly change the order of the timepoints for each gene) or vertically (shuffle at random data for a gene at a given timepoint for the same timepoint of another gene), or both. Using any of these methods we were never able to identify more than 15 false positives by Fourier analysis on the randomized data, using the same arbitrary cut-off that identified 810 genes with the unrandomized data. Hence the number of false positives due to noise alone is very low (<2%). False positives probably only arise in our data because of the arbitrary nature of our cut-off. We are currently using our knowledge of yeast biology in terms of function of the positives, and identifiable promoter elements within the upstream regions of the positives, to define more reasonably our cut-off. In addition, for genes whose transcripts peak late in G1, we observe a positive correlation (~0.4) between the inducibility of a transcript by the cyclin Cln3, and the peak to trough ratio of those genes. In contrast for the same genes we observe a negative correlation (~-0.4) between the inducibility of a transcript by Clb2 overexpression and their peak to trough ratios. This suggests that the transcriptional induction or repression of a gene in response to cyclin overexpression is physiologically relevant in that it mirrors the *in vivo*situation. It also tells us that genes which peak late in G1 are generally induced by Cln3 and repressed by Clb2. It has been previously noted that genes with SCB elements can be repressed by Clb2 (16), but it was observed that Clb2 was not necessary for repression of MCB regulated genes. We find that Clb2 does indeed repress genes with either SCB elements or MCB elements. Of genes that we identified as cell cycle regulated, and induced by Cln3, 122 were repressed by Clb2, 36 of which have both SCBs and MCBs, 11 of which have only SCBs, and 58 of which have only MCBs. This is compelling evidence that Clb2 can and does repress genes that peak in late G1 due to the presence of MCB elements in their promoters. Conversely I found that many genes whose transcription peaks in

M-phase are induced by Clb2 overexpression, and about half of these are repressed by Cln3 overexpression. Many of these Cln3 repressed genes are involved in mating. It has been previously noted that as cell pass the commitment point that they become refractory to arrest by mating pheromones (18). Transcriptional repression of genes required for the mating response by G1 cyclins might form part of that mechanism.

The major functions of the cell cycle regulated genes are DNA replication, DNA repair, budding, glycosylation, secretion, transcription, chromatin maintenance and organization, nuclear division and mitosis, proteolysis, and mating. It is instructive to look at the pattern of expression of genes involved in a particular process. For instance, we can trace the expression of many genes somehow involved in DNA replication. In the G1 cluster are 23 genes with known functions in DNA replication. These genes include subunits of the DNA polymerases and their accessory factors (e.g. CDC2, POL1, POL2, POL 12, POL32, PRI2, RFA1, RFA2, RFC3, RFC5), 5 genes involved in nucleotide synthesis (e.g. RNR1, RNR3, CDC21) and two genes involved in initiation of replication (CDC7 and CDC45). Many genes involved in DNA repair, especially mismatch repair, achieve peak expression in G1 phase. These include PMS1, MSH6, MSH2, RAD53, KIM2, RAD51, RDH51 RAD27, DUN1, OGG1, RAD5, and RHC18. It would seem from this that repair of DNA lesions may be a normal part of S-phase. Also at the same time, proteins involved in chromatin structure and chromatid cohesion are expressed, including CAC2, MCD1, ASF1 and 2, SMC1 and 3, MIF2, CIN2, SAS2, CSE4, and HFI1. Later, in the S-phase cluster, when S-phase is actually occurring, only a few proteins directly involved in DNA synthesis reach peak expression. These include RFA3, RFC1, and ORC3. Also reaching peak expression are the chromatin proteins PDS1, SWI3, and TBF1, and of course the nine histone or histone-like proteins HHF2, HTB1, HHT1, HHT2, HHO1, HTA1, HHF12, HTB2, HTA2. The

histone genes are the most tightly co-regulated set of genes we have seen, and the oscillation in their expression has one of the greatest amplitudes (about 20-fold in our experiments). By S/G2 the only replication gene peaking is that for ORC1, the largest subunit of the origin recognition complex. Perhaps extra Orc1 is needed for the new origins now present in the G2 nucleus. In G2/M the MCM gene Cdc47 reaches its peak, presumably to help set up origins for the next cell cycle, and in M/G1 the MCM genes CDC46 and CDC54 reach their peaks. Also reaching a peak at this time are CDC6 (though we did not find this gene in our analysis; see above) and the two repair genes RAD2 and HDF2 (a KU 80 homolog). Thus many genes needed for replication and repair reach peak expression just before they are needed, the histones peak exactly at the time they are needed, and a few genes important for initiation peak well in advance of the next round of S-phase (ORC1, CDC47, CDC46, CDC54, CDC6) presumably to set up initiation complexes. Only three known initiator genes, CDC7, CDC45, and DBF4 (which we did not find in our analysis) peak just before S-phase, suggesting that these may be particularly important to the trigger for replication.

Conclusions:

I have identified several good candidate Cdc28 substrates by use of a novel substrate-trapping mutant of Cdc28 in various 2-hybrid screens. Two of these candidates, Zds1 and Ydr130c, are particularly intriguing because they only give a detectable 2-hybrid interaction with the mutant Cdc28, and not the wild type. Zds1 has already been tied genetically to Cdc28, which makes its interaction with Cdc28 particularly compelling, whereas Ydr130c is a novel protein, for which no information is readily available. This provides for several obvious experiments, such as basic characterization of the gene and its product, including deletion analysis and epitope

tagging of the protein to investigate whether it can co-precipitate with mutant Cdc28, or if it is a phosphoprotein. For the 2-hybrid part of my project I will concentrate on these two positives, and consider sharing the information pertaining to the other positives with other interested parties for collaborative considerations. I feel that this project is very strong, and well poised to make some important breakthroughs, which will garner attention from throughout the cell cycle and yeast fields.

The genomic microarray projects that I have been involved with have generated enormous amounts of high quality data, which we will further analyze. A high priority for us is to investigate means by which this mass of data can be publicly available in the most easily accessible and readable form. I have spent considerable time developing a display mechanism for the data in the form of an Excel spreadsheet that we will make available for downloading via the Web. This document contains high quality information for each gene that we have identified as cell cycle regulated, including a functional description, whether it is induced or repressed by either Cln3 or Clb2, and the sequence and location of putative promoter elements for every gene in which we have found them. This project is only 8 months old, and I am already co-first authoring a paper detailing our findings, which we intend to submit to the distinguished journal Science. This represents a truly phenomenal rate of discovery, and demonstrates the power of using such genomic technologies. I intend to analyze further our dataset, and to address in greater detail the problem of potential false positives.

References:

1. P. Nurse, *Nature* **344**, 503-8 (1990).
2. D. O. Morgan, *Nature* **374**, 131-134 (1995).
3. P. D. Jeffrey, et al., *Nature* **376**, 313-320 (1995).
4. F. H. Espinoza, J. Ogas, I. Herskowitz, D. O. Morgan, *Science* **266**, 1388-1391 (1994).
5. V. Measday, et al., *Mol Cell Biol* **17**, 1212-23 (1997).

6. V. Measday, L. Moore, J. Ogas, M. Tyers, B. Andrews, *Science* **266**, 1391-1395 (1994).
7. C. Wittenberg, K. Sugimoto, S. I. Reed, *Cell* **62**, 225-37 (1990).
8. M. Tyers, G. Tokiwa, R. Nash, B. Futcher, *Embo J* **11**, 1773-84 (1992).
9. M. Tyers, G. Tokiwa, B. Futcher, *Embo J* **12**, 1955-68 (1993).
10. C. Kuhne, P. Linder, *EMBO J* **12**, 3437-3447 (1993).
11. C. B. Epstein, F. R. Cross, *Genes Dev* **6**, 1695-706 (1992).
12. E. Schwob, K. Nasmyth, *Genes Dev* **7**, 1160-75 (1993).
13. I. Fitch, et al., *Mol Biol Cell* **3**, 805-18 (1992).
14. L. Dirick, T. Bohm, K. Nasmyth, *Embo J* **14**, 4803-4813 (1995).
15. D. Stuart, C. Wittenberg, *Genes Dev* **9**, 2780-2794 (1995).
16. A. Amon, M. Tyers, B. Futcher, K. Nasmyth, *Cell* **74**, 993-1007 (1993).
17. B. L. Schneider, Q.-Y. Yang, A. B. Futcher, *Science* **272**, 560-562 (1996).
18. L. J. Oehlen, F. R. Cross, *Genes Dev* **8**, 1058-1070 (1994).
19. A. Amon, S. Irniger, K. Nasmyth, *Cell* **77**, 1037-1050 (1994).
20. K. Nasmyth, *Curr Opin Cell Biol* **5**, 166-79 (1993).
21. M. Tyers, B. Futcher, *Mol Cell Biol* **13**, 5659-69 (1993).
22. K. N. Huang, S. A. Odinsky, F. R. Cross, *Mol Cell Biol* **17**, 4654-66 (1997).
23. K. Levine, K. Huang, F. R. Cross, *Mol Cell Biol* **16**, 6794-803 (1996).
24. A. J. Flint, T. Tiganis, D. Barford, N. K. Tonks, *Proc Natl Acad Sci U S A* **94**, 1680-5 (1997).
25. P. James, J. Halladay, E. A. Craig, *Genetics* **144**, 1425-36 (1996).
26. Y. Yu, et al., *Mol Cell Biol* **16**, 5254-63 (1996).
27. M. Mizunuma, D. Hirata, K. Miyahara, E. Tsuchiya, T. Miyakawa, *Nature* **392**, 303-6 (1998).
28. M. M. Luke, et al., *Mol Cell Biol* **16**, 2744-55 (1996).
29. D. A. Lashkari, et al., *Proc Natl Acad Sci U S A* **94**, 13057-62 (1997).
30. J. L. DeRisi, V. R. Iyer, P. O. Brown, *Science* **278**, 680-6 (1997).

Appendices:

Curriculum vitae:

Personal Details

PII Redacted

Full name :	Gavin James Sherlock
[redacted]	[redacted]
Telephone :	(516) 367 8333
Fax :	(516) 367 8369
email :	sherlock@cshl.org

Work Experience and Training

1990:	Three months full-time research for Amersham International Plc.
1990-91:	Two term research project at the Regional Molecular Genetics Laboratories, St. Mary's hospital, Manchester.
1991-94:	Ph.D. project investigating the cell cycle of *C. albicans* in John Rosamond's laboratory.

As an undergraduate I undertook a 3 month vacation project at the laboratories of Amersham International, characterising a phage of the pathogenic bacterium *Campylobacter jejuni*. Whilst in the 3 months I did not have enough time to complete the project, I was able to learn several laboratory techniques, and understand how an industrial research laboratory is run.

As part of my Bachelor's degree I carried out a two-term research project in the Regional Molecular Genetics laboratories at St. Mary's Hospital in Manchester. The aim of the project was to isolate sequences from human chromosome 2q which have been conserved throughout evolution. I isolated and partially characterised two sequences. Again this project introduced me to new techniques, and a considerable report was submitted in support of my degree.

My doctoral work focused on control of the cell cycle of the human fungal pathogen, *Candida albicans*, which can cause fatal infection in immunocompromized patients, such as those with cancer or AIDS. I worked for three years in John Rosamond's laboratory using the yeast *Saccharomyces cerevisiae* as a surrogate genetic system to isolate *C. albicans* cell cycle genes by complementation of conditional mutants. In 1992, after submission of my transfer report, my introduction was considered good enough for me to be commissioned to write a review article on the control of START in *S. cerevisiae*. I cloned and characterised cyclins and a Cdk from *C. albicans*, and first authored a paper detailing these results. I also set up a system for inducing and monitoring synchronous cell division and hyphal formation of *C. albicans*. In addition I have been commissioned to write a further review on the control of the *S. cerevisiae* cell cycle, based on the introduction to my Ph.D. thesis, due to a recommendation from my examiners.

Since being at Cold Spring Harbor I have designed and made vectors to facilitate easy mutagenesis of yeast genes. I have used them to mutate the phosphorylation sites of Cdc6, Cin8 and Chs2 in various combinations. I have also been investigating initiation of DNA synthesis in phosphorylation site mutants of Orc2 and Orc6 in a collaboration with Joachim Li. These studies have recently seen fruition, and we will soon be preparing a paper detailing these results.

In summary I have extensive research experience which is particularly relevant to the proposed research. I have a good understanding of yeast cell biology and the cell cycle in particular. These attributes are ideal for the proposed research, which will allow me to use and extend my techniques, as well as augmenting my scientific repertoire with new techniques of protein chemistry.

Education and Qualifications

1988-1994 : Manchester University,
Oxford Road,
Manchester M13 9PT.
U.K..

1994: Ph.D.: Molecular analysis of G1 in *Candida albicans*. Ph.D. thesis, The Victoria University of Manchester, UK.

1991: Bachelor's Degree: 1st Class Honours Degree in Genetics.

1988: `A' levels: A grades in Biology, Chemistry, Physics and Mathematics.

Awards

1988 Awarded John Buckley Scholarship for Science as an undergraduate.
1991 Awarded a Wellcome Trust Prize Studentship
1996 Awarded Cold Spring Harbor Fellowship

Publications

Spellman, P.*, Sherlock, G.*, Zhang, M., Iyer, V., Brown, P., Botstein, D. and Futcher, B. The cell cycle regulated genes of yeast. *In preparation* (* = joint 1st authors).

Sherlock, G. and Futcher, B. Cdc6 has dual functions in the yeast mitotic cell cycle. *In preparation.*

Donovan, J., Sherlock, G., Futcher, B. and Johnston, L. A novel gene involved in the regulation of the exit from mitosis in budding yeast. *In preparation.*

Sherlock, G., Bahman, A.M., Mahal, A., Shieh, J.-C., Ferreira, M., and Rosamond, J. (1994). Molecular cloning and analysis of *CDC28* and cyclin homologues from the human fungal pathogen *Candida albicans*. *Mol. Gen. Genet.* **245**, 716-723.

Sherlock, G. and Rosamond, J. (1993). STARTing to Cycle: G1 Controls Regulating Cell Division in Budding Yeast. *Journal of General Microbiology* **139**, 2531-2541.

Abstracts

Spellman, P., Sherlock, G., Zhang, M., Iyer, V., Brown, P., Botstein, D. and Futcher, B. The cell cycle regulated genes of yeast.. FASEB yeast chromosome structure, replication and segregation meeting, Colorado, 1998.

Sherlock, G. and Futcher, B. Cdc6 has dual functions in the yeast mitotic cell cycle. Cold Spring Harbor Cell Cycle Meeting, 1998.

Spellman, P., Sherlock, G., Futcher, B., Brown, P.O., and Botstein, D. Identification of RNAs transcribed in a cell cycle dependent manner in the yeast *Saccharomyces cerevisiae*. Cold Spring Harbor Cell Cycle Meeting, 1998.

Donovan, J., Sherlock, G., Futcher, B. and Johnston, L. A novel gene involved in the regulation of the exit from mitosis in budding yeast. Cold Spring Harbor Cell Cycle Meeting, 1998.

Futcher, B., Yang, Q.-H., Sherlock, G., Marshak, D. and Schneider, B. SIC1 and other substrates of CDK kinases. Dept. of Defense Breast Cancer Research Program Meeting: Era of Hope; Maryland, 1997.

Sherlock, G., Schneider B. and Futcher , B. Epitope tagging of yeast proteins using PCR approaches. FASEB yeast chromosome structure, replication and segregation meeting, Colorado, 1996.

Rosamond, J., Shieh, J.-C. and Sherlock, G. Control of cell division and development in *Candida albicans*. Keystone Meeting on Host-fungal pathogen interactions, Taos. *Journal of Cellular Biochemistry* **19S**, 153 (1995).

Sherlock, G., Shieh, J.-C., Ball, T., Bahman, M. and Rosamond, J. Analysis of G1 functions in *Candida albicans*. Cell Cycle Meeting, Cold Spring Harbor Laboratories, 1994

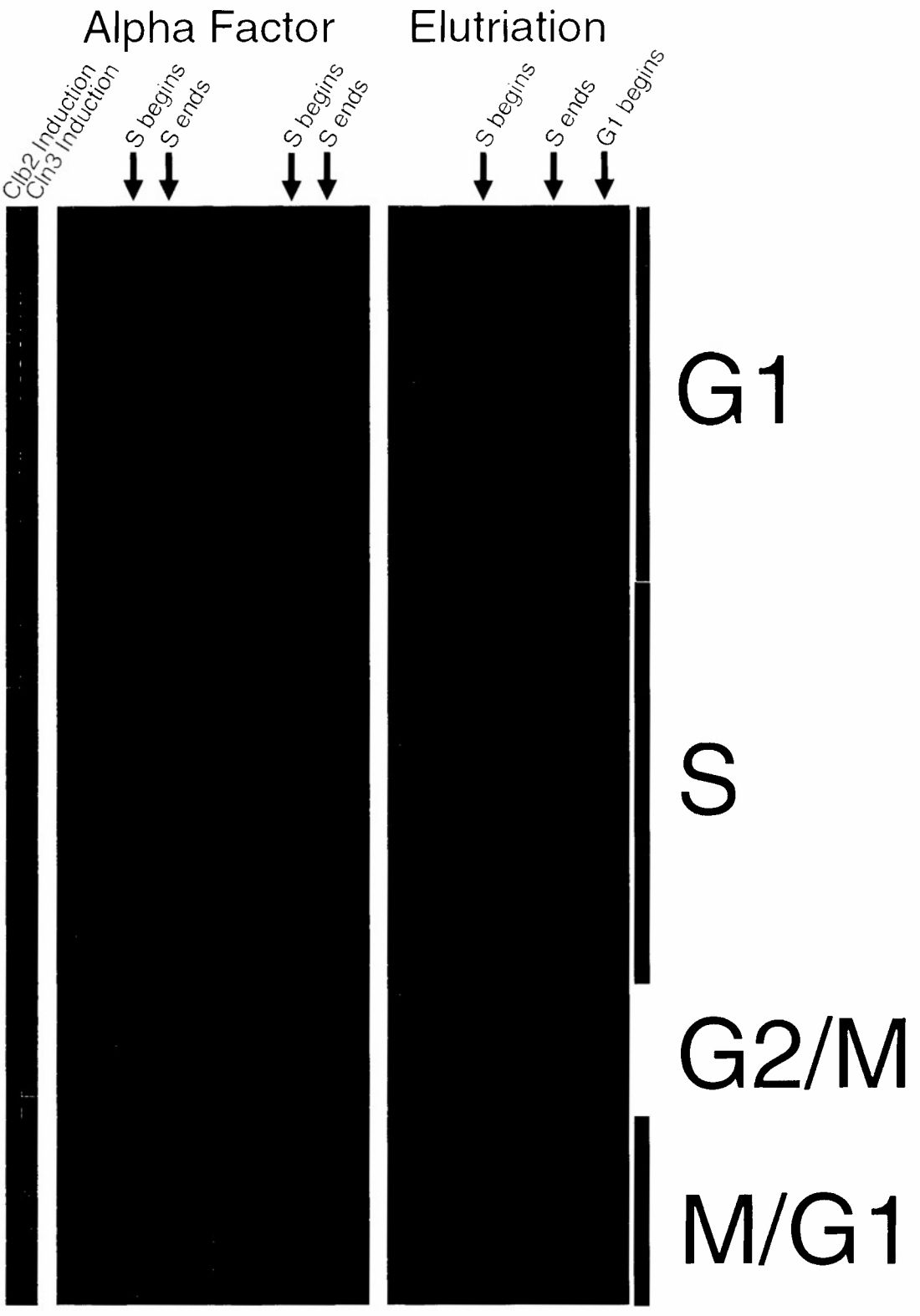
Sherlock, G., Shieh, J.-C., Ball, T. and Rosamond, J. Analysis of G1 functions in *Candida albicans*. British Yeast Meeting, University of Sheffield, 1994.

Sherlock, G., Shieh, J.-C. and Rosamond, J. Analysis of G1 in *Candida albicans*. Predoctoral Meeting at Liverpool University, 1993. Prize awarded for best oral presentation.

Sherlock, G., Shieh, J.-C. and Rosamond, J. Analysis of G1 functions in *Candida albicans*. Abstracts of the 16th Yeast Genetics & Molecular Biology Meeting, Madison, WI, USA; p.57.

Figure Legend:

The attached color figure shows a graphic represenatation of all the genes that we have determined as being cell cycle regulated. Green represents a transcript level lower than an asychronous population, and red a higher transcript level.


Alpha Factor
Elutriation
Clb2 Induction
Cln3 Induction
S begins
S ends
S begins
S ends
S begins
S ends
G1 begins
G1
S
G2/M
M/G1